SPECIFICATION

Title: Preferential Display

Background of Invention

[0001] Methods for assaying gene expression can be classified into two major types: open methods, which do not require prior knowledge of the genes being measured, and closed methods, which measure expression levels of already collected clones or sequences. Some expression analysis techniques can only measure on a gene-by-gene basis while others can assay multiple genes simultaneously. Finally, some methods can directly measure differential expression between two samples and some examine expression levels from one sample at a time, followed by computation based comparisons. Understanding differences between these methods is essential for choosing the best technology for a given application. Regardless of the methods chosen researchers must identify or access through databases vast quantities of expression information to find the actual cause and effect on the gene expression (_Ref488664383Lennon, G.G. (2000) DDT, 5:59-66_Ref488664383).

[0002] The history of gene expression analysis began when laboratory methods were developed to examine expression of individually known genes. The northern blot technique, introduced in 1977, hybridizes labeled DNA or RNA prob_Hlt488664282_Ref488664328es of known genes to RNA blots.(Alwine, J.C. et al. (1977) Proc. Natl. Acad. Sci. U. S. A. 74: 53505354_Hlt488664282_Ref488664328). The resulting expression patterns of mRNA transcripts were then read. This technique is still widely used to confirm the results of other types of gene expression studies. In 1977, another method was published that protects a DNA-labeled probe against degradation by the single-stranded nuclease S1 if the probe is annealed to an RNA (Berk, A.J. and Sharp, P.A. (1977) Cell 12:721732). Ten years later, RNase protection assays were developed to detect the expression of specific, previously characterized RNAs and to compare their levels of expression(Berk, A.J. and Sharp, P.A. (1977) Cell 12:721732). With this technique, a specific labeled cDNA forms a hybrid with its corresponding mRNA. When exposed to a single-strand-specific nuclease, the hybrids resist degradation and can be detected using gel electrophoresis. A later approach, differential plaque-filter hybridization, can detect differences in the expression of cloned cDNAs between two samples (_Ref488664612Lee, J.J. and Costlow, N.A. (1987) Methods Enzymol. 152:633648_Ref488664612).

[0003] In 1993, subtractive hybridization techniques became available for constructing subtractive cDNA libraries. This methodology hybridizes cDNA from one pool to mRNA from another (_Ref488664642Hedrick, S.M. et al. (1984) Nature 308:149153_Ref488664642). Then, cDNA libraries are constructed from the transcripts that are not hybridized, these being used to identify specific mRNAs (_Ref488664659Swaroop, A. et al. (1991) Nucleic Acids Res. 25:1954_Ref488664671_Ref488664659; Lisitsyn, N. et al. (1993) Science 259:946951_Ref488664671). A modification of this technique, representational difference analysis (RDA), also uses preferential amplification of non-subtracted fragments. In RDA, "representations" or simplified versions of the genomes being studied (amplicons) are created using restriction digestion. This method was first developed to examine the differences between genomes, but has proven useful for cloning differentially expressed genes. From this method, suppressive subtractive hybridization (SSH) was derived, which enables further suppression amplification of non-subtracted fragments (Lisitsyn, N. et al. (1993) Science 259:946951). SSH combines normalization (equalizing the abundance of cDNAs within the target population) and subtraction (excluding the common sequences between the target and driver populations) in a single procedure. Results from both RDA and SSH should be validated using other methods.

[0004] Early gene expression methods, such as those already mentioned, are relatively small-scale techniques. They either focus on measuring mRNA expression levels for individual, well-characterized genes, or use in

vitro nuclear "run-on" transcription assays to determine the transcriptional profiles of several active genes simultaneously (_Ref489091994Greenberg, M.E. and Ziff, E.B. (1984) Nature 311:433438_Ref489091994; _Ref489092005Marzluff, W.F. (1978) Methods Cell Biol. 19:317331; Manley, J.L. and Gefter, M.L. (1981) Anal. 2:369382_Ref489092005). They are therefore inadequate for conducting large-scale screening and developing expression profile patterns for tissues or cells (the basic requirements for efficient pharmaceutical research). Thus, several newer methods for high-throughput screening (HTS) have been developed over the past decade, including differential display, expressed sequence tag (EST) methodology and many array techniques. Collectively, they have made it possible to identify the expression levels of novel genes and characterize them, correlate mRNA expression patterns in many tissue types with disease states, identify side effects of current and experimental treatments, and determine the effects of compounds on non-target tissues.

[0005] Differential display of eukaryotic mRNA, first reported in 1992, was a major advance in the comparison of gene expression differences between cells or tissues (_Ref489091961Liang, P. and Pardee, A.B. (1992) Science 257:967997_Ref489091961). Encompassing the use of either arbitrarily or specifically primed PCR, it is perhaps the most widely used method involving gel electrophoresis for comparing gene expression. Both methods amplify partial cDNAs from subsets of mRNA samples by using reverse transcription and PCR. These short cDNA fragments are then typically displayed on polyacryl-amide gels. Differential display can simultaneously measure both up- and down-regulation across tens of samples.

[0006] Originally, this method used an oligo(dT) primer with an anchor of one or two bases at the 39 terminal. Reverse transcription and denaturation were followed by arbitrary priming on the resulting first strand of cDNA. A series of products were then derived from the 39 end of the mRNAs by using PCR with the original primer (a radiolabeled nucleotide) and a set of short, random decamer primers. Each random primer annealed to the mRNA at a different position relative to the anchor primer. Products showing significant differential expression were sequenced after size fractionation of the PCR sample using denaturing gel electrophoresis, generally after overnight autoradiographic exposure.

[0007] EST methodology can determine the expression profile of an entire cell or tissue under analysis. During the 1990s, EST methodology played the largest role in increasing the catalog of known genes. Using this approach, cDNA clones are randomly picked and a single pass of sequencing is performed from one or both ends of each clone. Subsequent comparison with existing sequence databases immediately identifies novel sequences. Measuring how often a given sequence appears in a (representational) library enables the estimation of expression levels for each gene.

[0008] Although this method can accurately identify the presence of a proportion of genes, relatively low sampling (typically 5,00010,000 sequences are generated from a tissue containing >20,000 distinct transcript types) makes it difficult to measure abundance of expression or to identify differentially expressed genes except where genes are highly up- or down-regulated.

[0009] Serial analysis of gene expression (SAGE) can potentially tag and analyze all transcripts in a given cell population or tissue. It has been used to successfully compare expression profiles between normal and cancerous cells, and detect p53 levels prior to apoptosis (_Ref489091855Zhang, L. et al. (1997) Science 276:12681272_Ref489091867_Ref489091855; Polyak, K. et al. (1997) Nature 389:300305_Ref489091867). In theory, SAGE is an "open" system. However, in practice, the short length of the tags means that it is most useful for expression profiling of fully sequenced genes. Thus, the value of this technique might increase as the Human Genome Project progresses.

[0010] This method uses two samples that are ligated and tagged with separate primers and then amplified. Subsequently, the primers are removed, revealing sticky ends that form concatemers. The concatemers are both cloned into a vector, with sequence information for the two different cDNA tags contained between anchoring sites. This cloning and sequencing process is time-consuming, as it must be performed for each sample and followed by extensive computational analysis.

[0011] The public EST efforts, spearheaded by sequencing work at Washington University (St. Louis, MO, USA) (Schena, M. et al. (1996) Proc. Natl. Acad. Sci. U. S. A. 93:1061410619) and the arraying efforts of the IMAGE Consortium (founded by researchers at the Lawrence Livermore National Laboratory, Columbia University, National Institutes of Health and Centre National de la Recherche Scientifique) (_Ref489091835Lennon, G.G. et al. (1996) Genomics 33:151152_Ref489091835), have made sequences and clones for more than one million cDNA clones publicly available. A network of five distributors (see http://wwwbio.llnl.gov/bbrp/image/image.html) across the globe supplies researchers with clones and related research services, such as sets of sequence-verified cDNA clones spotted onto nylon membranes. As standard laboratory protocols can be used and the filters are commercially available at a relatively modest cost, they are a popular forerunner to microarrays. Hybridization of radioactively labeled complex RNA (_Ref489091805Gress, T. et al. (1992) Genome 3:609619_Ref489091805) to these membranes yields signals for moderately and abundantly expressed genes and, depending on several factors, some of the less abundant transcripts. Thus, differential expression is best measured using genes that are moderately expressed in at least one of the two (or more) states under study.

[0012] DNA microarrays measure expression by using templates containing hundreds or thousands of probes that are exposed simultaneously to a target sample. They make it possible to systematically survey DNA and RNA variation for the first time and are becoming a standard tool for drug discovery and evaluation. Microarray techniques are so powerful that their uses are often limited largely by the challenge of managing and analyzing the data they generate. DNA microarray technology evolved from a paper published in 1975 by E.M. Southern (the originator of the Southern blot), who showed how a solid support could be used to examine nucleic acids (_Ref489091774Southern, E.M. (1975) J. Mol. Biol. 98:503517_Ref489091774). This was advanced by the development of non-porous solid supports, leading to miniaturization and the use of fluorescence-based detection methods. The two main types of templates are long DNA fragments (_Ref489091723Schena, M. et al. (1995) Science 270:467470_Ref489091723) (over 100 base pairs) and oligo-nucleotides (generally 1825 mers). Microarrays are expensive, although efficiencies should improve and costs should drop dramatically in the next couple of years, enabling these tools to become accessible to most research laboratories. Besides cost, microarrays are limited by the fact that they can only probe genes for which clones or sequences are already available. Further-more, their accuracy can be limited by the purity of the RNA and the quantity of RNA for each hybridization.

[0013] By understanding gene expression patterns, researchers can gain information that can link sites of expression, bio-chemical pathways, and normal or pathological functions in organs and whole organisms. Because of their speed and breadth, microarrays should impact genetic profiling in several ways: Accelerate the understanding of the molecular basis of disease or environmental stresses, Improve knowledge of model systems, Explore pathogens, pathogenic, environmental (microgravity) reactions in terms of gene expression, Pinpoint new molecular level explanations to environmental effects, and Examine efficacy and toxicity responses to environmental or other external simulates.

[0014] Microarrays have already determined how several important genes are abnormally regulated in disease. For example, a microarray of approximately 100 genes that have a role in inflammation was used to examine rheumatoid tissue. This revealed upregulation of the genes encoding interleukin-6 and several matrix metalloproteinases (Schena, M. et al. (1995) Science 270:467470). In another instance, a novel gene involved in promoting tumors was discovered by using a 1000-element micro-array of unknown cDNAs to examine how treatment with phorbol testers affects expression levels (_Ref489091711Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. U. S. A. 94:21502155_Ref489091711). Microarrays should provide more detailed knowledge about pathogens by systematically examining every gene in a microbe to uncover the overall expression pattern. In addition, microarrays will continue to contribute to the understanding of responses to drug treatments. For example, a recent study used microarrays to measure the effects of kinase inhibitors on the entire yeast genome by measuring changes in mRNA levels before and after treatment (_Ref489091691Gray, N.S. et al. (1998) Science 281:533538_Ref489091691). In another example, microarray studies of yeast cells showed that the immunosuppressive drug FK506 had the same effect on gene expression level patterns as ablation of the gene that FK506 suppresses. Furthermore, this study showed that, in the absence of this gene, FK506 affected expression levels in other ways

(_Ref489091674Marton, M.J. et al. (1998) Nat. Med. 4:1293130_Ref489091674). This suggests that the drug might have more than one target. Microarrays are also proving useful in the determination of drug toxicity (_Ref489091660Braxton, S. and Bedilion, T. (1998) Curr. Opin. Biotechnol. 9:643649_Ref489091660).

[0015] Expression profiling using cDNA microarrays begins by arraying many gene-specific amplicons derived from the cDNA clones onto a single matrix. Using two-color hybridization, cDNA representations of total RNA pools are created from test and reference cells, fluorescently tagged with two different colors, then mixed together before being hybridized to the matrix. For each transcript, the resulting fluorescence signals reflect the difference in abundance between the two samples. Two-color hybridizations provide rapid comparisons between the two samples, but they do not measure the absolute levels of gene expression for either sample. By contrast, one-color hybridization is slightly slower, as hybridizations of the two samples must be performed separately to reach meaningful comparisons. However, each one-color hybridization measures absolute levels of gene expression rather than comparative levels. After these actual levels are recorded in databases, they can be compared with levels from other samples without the need to perform comparative experiments. Although performing 1000 two-color hybridizations results in 1000 pair-wise comparisons, conducting 1000 one-color hybridizations yields almost half-a-million pair-wise comparisons, as the absolute values of one-color hybridizations can be evaluated against each other.

[0016] Using either the one- or two-color methods, microarray experiments must be performed repeatedly to ensure accuracy of the data. However, computational averaging of the signals of one-color hybridizations from multiple independent samples is more straightforward. The choice between using one-color versus two-color methods depends on several factors, including the number of transcripts under examination, the need for speedy results and cost differences. Hence, one-color hybridizations are often more useful for surveying a large number of genes, while two-color hybridizations can be preferable for more restricted sampling. Major portions of this background information was taken from the Lemon et.al. reference (Lennon, G.G. (2000) DDT, 5:59-66).

Brief Description of Drawings

[0017] FIG. 1 is a schematic diagram illustrating the preferential display technique. The figure shows the end products of unique RNA and cDNA strands isolated after degradation of the cDNA/RNA compliments. FIG 2 is a schematic diagram of the final treatment with RNase to degrade the single stranded RNA's present in each sample and the PCR amplification of the isolated cDNA which are uniquely expressed in each sample.

Detailed Description

[0018] The major problem in present gene expression technology is that too many genes are identified thus causing much wasted time and cost in sequencing redundant sequences found in both normal and diseased or diseased tissues.. The method is a revolutionary differential gene display technology that provides a simplified, cost-effective method to efficiently identify and accurately isolate validated gene sequences between normal and diseased or diseased states of tissues. The method utilizes a combination of biomolecular chemistry methods to eliminate redundant sequences and fluorescence dye assay to identify unique sequences found in normal and/or diseased tissues. The method can be implemented in a single reaction tube, is amendable to miniaturization, and is extremely cost-effective which would be of benefit in toxicogenomic and drug discovery applications. It is estimated redundant sequences amounts to 98% of the total of all genes expressed in current display technologies.

[0019] The invention is a method called Preferential Display. The approach starts with the sample collection and categorizing of the cells, tissues or blood samples. Once the normalized control cells are isolated from the diseased state expressed RNA's are isolated using standard methods. The RNA's are then placed into two tubes of normal and the other diseased RNA's for a total of four tubes. One normal and diseased are then RT-PCR with dye labeled oligos producing fluorescence labelled cDNA's in each tube. At this point of the reaction, RNA's of normal and diseased are added to their complementary tubes, normalized cDNAs with diseased RNAs and diseased cDNAs with normalized RNAs. Common sequences in each tube hybridize to form cDNA/RNA compliments. The tubes are then treated with Exonuclease III or VII to degrade all the hybridized complements. RNase is then added to digest the single stranded RNA's in each tube. The remaining undigested cDNA are unique dye labeled sequences expressed in either normal or diseased states. Additional PCR can be ran to increase the cDNA present in each tube before running on a gel or high throughput sequencer.

[0020] The preferred embodiment of this invention begins with the Total RNA isolation and purification from cells/tissue using Totally RNA, RNA extraction kit from Ambion (Austin, TX). Reactions are carried out as per manufacturer"s instructions.

[0021] RNA content is determined by adding 5 _l sample to 495 ul TE buffer and absorbance read at 260 nm. Concentration of RNA in _g/_l is determined by dividing absorbance at 260 nm by 0.025x (dilution factor). The average yield of RNA was 0.3 _g/_l. The major components that make up this invention includes the following materials.

Materials:
·dH$_2$O
·10X PCR Buffer (100mM Tris-Cl, pH 8.4, 500mM KCl, 15mM, 15mM MgCl$_2$ and 0.01% gelatin), GeneHunter Corporation™, Nashville, TN
·5X RT Buffer (125mM Tris-Cl, pH 8.3, 188mM KCl, 7.5mM MgCl$_2$ and 25m m DTT), GeneHunter Corporation™, Nashville, TN
·dNTP (250μM) and (25μM) GeneHunter Corporation™, Nashville, TN
·H-T$_{11}$M, where M could be G, A or C (2μM), GeneHunter Corporation™, Nashville, TN
·H-APX, where X could be any arbitrary primer from 1 to 80 (2μM), GeneHunter Corporation-™, Nashville, TN
·MMLV Reverse Transcriptase (100U/μl), GeneHunter Corporation™
·RNAse H (60U/μl), diluted to 4U/μl with TE buffer, Takara, Shiga, Japan
·Taq DNA Polymerase (5U/μl) Qiagen™, Valencia, CA
·Taq Polymerase Master Mix, Qiagen™, Valencia, CA
·Exonuclease III (65U/μl), GibcoBRL™ Life Technologies™, Carlsbad, CA
·Exonuclease VII (152U/μl), GibcoBRL™ Life Technologies™, Carlsbad, CA
·P^{33}_ATP (250μCi), NEN® Life Science Products, Boston, MA
·cDNA (0.01μg/μl)
·RNA (0.1μg/μl)

[0022] For the best mode of embodiment of the invention, the first step is the reverse transcription of mRNA from a total RNA pool. The Total RNA pool is obtained from mice heart and brain tissues provided by Clontech™, Palo Alto, CA. Each preferential display reaction is preceded by an RNAse H incubation to eliminate the mRNA from the cDNA/mRNA complex that is obtained from the initial RT reaction. Once this is accomplished, the cDNA is ready to hybridize with the foreign mRNA that is introduced and that has some portions complementary to this cDNA.

[0023] Example 1Reverse Transcription ProtocolTh e Reverse Transcription protocol was obtained from the RNAimage kit manufactured by the GenHunter Corporation™, Nashville, TN.

Component	Amount
dH₂O	9.4
5X RT Buffer	shapeType20fFlipH0fFlipV0shapePath4fFillOK 0fFilled0lineColor16777215fArrowheadsOK1fL ine1fLayoutInCell14.0
dNTP (250μM)	1.6
H-T₁₁M (2μM)	2.0μl
Total RNA (0.1μg/μl)	2.0μl
MMLV RT	1.0μl

Thermocycling Conditions for the RT Reaction
65°C for 5 minutes
37°C for 10 minutes
*Pause: Add MMLV RT Enzyme
37°C for 50 minutes
75°C for 5 minutes
Hold at 4°C

[0024] Example 2

[0025] RNAase H Protocol

[0026] Following the RT reaction from Example 1, an RNAse H incubation is performed by adding 1.0μl of RNAse H (4U/μl) to the 20μl RT Reaction product:

Component	shapeType20fFlipH0fFlipV0shapePath4fFillOK 0fFilled0fArrowheadsOK1fLayoutInCell1Amou nt
dH₂O	9.4
5X RT Buffer	4.0
dNTP (250μM)	shapeType20fFlipH0fFlipV1shapePath4fFillOK 0fFilled0fArrowheadsOK1fLayoutInCell11.6
H-T₁₁M (2μM)	2.0μl
cDNA/mRNA complex (0.01μg/μl)	From 2.0μl of total RNA (0.1μg/μl)
MMLV RT	1.0μl
RNAse H	shapeType20fFlipH0fFlipV0shapePath4fFillOK 0fFilled0lineEndArrowhead1fArrowheadsOK1f LayoutInCell11.0μl

[0027] At this point the foreign cDNA is ready to be mixed with the foreign RNA.. For example, if brain cDNA is obtained, then heart RNA can be introduced and vice versa. Two different additional embodiments are used for this step depending on the enzyme utilized. The first embodiment is the Exonuclease III (digests dsDNA and portions of the DNA/RNA complex) embodiment and the other embodiment is the Exonuclease VII (digests ssDNA).

[0028] Example 3

[0029] Exonuclease III Approach

Component	Amount
dH$_2$O	4.0µl
10X PCR Buffer	0.8µl
cDNA (RT)	2.0µl
RNA	2.0µl
Exonuclease III	2.0µl

Thermocycling Conditions for the Exonuclease III Reaction
95°C for 10 seconds
*Pause: Add RNA from foreign tissue
60°C for 3 minutes
37°C for 10 seconds
*Pause: Add Exonuclease III
37°C for 30 minutes
95°C for 10 minutes
Hold at 4.0°C

[0030] Example 4

[0031] Exonuclease VII Approach

Component	Amount
dH$_2$O	3.0µl
10x PCR Buffer	1.0µl
cDNA (RT)	2.0µl
RNA	1.0µl
Exonuclease VII	1.0µl
Rnase H (4U/µl)	0.5µl
Taq Polymerase Master Mix	13.0µl
H-APX	2.0µl

Thermocycling Conditions for Exonuclease VII Approach
95°C for 10seconds
*Pause: Add RNA
60° for 3 minutes
*Pause: Add H-APX and Taq Poly Master Mix
40°C for 1 minute
30°C for 10 seconds
*Pause: Add RNAse H
30°C for 20 minutes
*Pause: Add Exonuclease VII
37°C for 30 minutes
95°C for 5 minutes
Hold at 4.0°C

[0032] Regardless of the approach used, the product of the preferential display reactions is followed by a PCR with the following conditions:

[0033] Example 5

[0034] PCR Theromcycling Protocol

Thermocycling Conditions for PCR Reaction
94°C for 3 minutes
94°C for 30 seconds
40°C for 2 minutes
72°C for 30 seconds
72°C for 5 minutes
Hold at 4°C

[0035] For the Exonuclease VII approach $0.2\mu l$ of P^{33}_ATP, $2.0\mu l$ of H-APX and H-T_{11}M are added to the preferential display product prior to the PCR reaction.

[0036] The Exonuclease III approach does not require these components, instead $2.0\mu l$ of H-APX, $2.0\mu l$ of H-T_{11}^-M, $0.2\mu l$ of P33_ATP, $1.6\mu l$ of dNTP ($25\mu M$), $1.2\mu l$ of 10X PCR Buffer and $2.2\mu l$ of dH$_2$O are added. These make a $20\mu l$ total reaction volume for PCR.

[0037] Genes display is carried out using QuickPoint™ Gel system (Novex; San Diego, CA). Assay procedure is same as recommended by the manufacturer. Briefly, the precast gel is pre-electrophoresed for 5 minutes. 20 _l of PCR reaction mixed with an equal volume of QuickPoint™ sample loading buffer is then heated to 80°C for 2 minutes and 1 _l is loaded immediately onto the gel (6% polyacrylamide). 0.5 _l of the sample loading buffer is loaded onto the remaining wells to assure a straight banding pattern. Electrophoresis is carried out in QuickPoint™ cell at 1200 V for approximately 10 minutes depending upon the length of the sequences. Following electrophoresis, the glass cassette enclosing the gel is washed in water for 5 minutes on a low speed shaker, dried in an oven at 80°C for 20 minutes and exposed to X-ray film overnight to generate an autoradiogram.

[0038] The present invention has been described with reference to particular preferred embodiments; however, the scope of the invention is defined by the attached claims and should be construed to include reasonable equilvalents.